

Recueil, analyse & traitement de données : Le questionnaire

Mise à jour du 21 mars 2011



Dernière version des diapos disponible ici : concevoir et administrer un questionnaire

Source des images indiquées au-dessous ou en cliquant sur l'image

Cours distribué sous licence **Creative Commons,** selon les conditions suivantes : 

bachelet@bigfoot.com
École Centrale de Lille
Villeneuve d'Ascq - France

Image: source

Source des images indiquées au-dessous ou en cliquant sur l'image

VISIT…

LANZAROTE
Caliente.COM

Comment utiliser ce cours :

1. Mettre les diapos en format plein écran en cliquant sur 
2. Faire défiler l'animation en cliquant **sur** les diapositives

(attention : cliquer sur une image ou un lien ouvre la page web correspondante)



Le questionnaire

On choisit un questionnaire lorsqu'on :

- Connaît bien le sujet de l'étude
- Veut des résultats quantifiés
- Cherche à valider et généraliser les résultats
- A les moyens de mener une telle enquête
    - accéder à la population
    - recevoir suffisamment de réponses
    - capacité à traiter les données *(argent, temps, ressources)*

Construire un questionnaire : un processus aux étapes bien définies

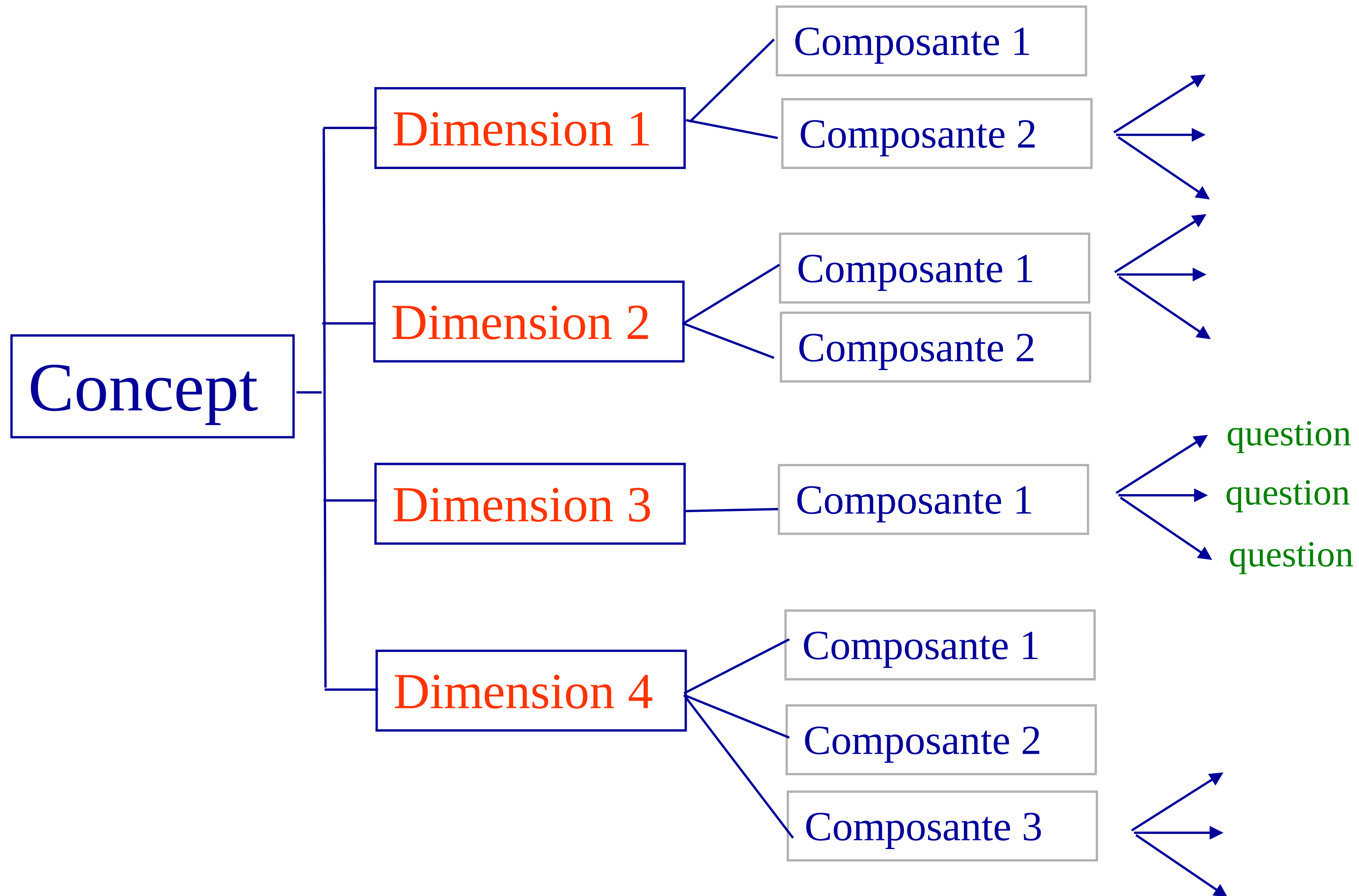
1. Fixer le type de questionnaire à employer, sa diffusion, la stratégie de collecte…
2. Rédiger une « version alpha » …
3. … que l’on testera sur un échantillon de personnes et/ou auprès d’experts
4. Pour obtenir une version finale et avoir un scénario de collecte parant aux risques prévisibles

… au total, le processus compte 9 étapes qui seront récapitulées plus bas

Lancer une campagne d’administration d’un questionnaire représente un investissement irréversible,
attention à ne pas se rater…

Utilisation ou copie interdites sans citation

Comment construire un questionnaire



Utilisation ou copie interdites sans citation

Questionnaire : cas de la croyance en dieu

Construisez un questionnaire pour étudier le concept de « croyance en dieu »

Utilisation ou copie interdites sans citation

Questionnaire : cas de la croyance en dieu

Dimensions	Composantes	Questions

Utilisation ou copie interdites sans citation

Questionnaire : cas de la croyance en dieu

Dimensions
Croyance
Expérientielle
Ritualiste
Conséquentielle

Utilisation ou copie interdites sans citation

Questionnaire : cas de la croyance en dieu

Dimensions	Composantes
Croyance	Croyance en dieu
	Croyance au diable
	Croyance à la trinité
	Croyance à l'enfer
Expérientielle	Apparitions
	Communication avec dieu
	Intervention de dieu dans sa vie
Ritualiste	Prière
	Messe
	Sacrements
	Pèlerinage
Conséquentielle	Pardonner
	Être honnête

Utilisation ou copie interdites sans citation

Questionnaire : cas de la croyance en dieu

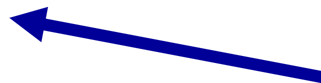
Dimensions	Composantes	Questions
Croyance	Croyance en dieu	Croyez-vous que dieu existe ?
		Comment vous représentez-vous dieu ?
	Croyance au diable	Croyez-vous à l'existence du diable ?
		S'agit-il d'un être qui pousse réellement à faire le mal ?
		N'est-il que la représentation abstraite du mal que subit l'humanité ?
	Croyance à la trinité	
	Croyance à l'enfer	
Expérientielle	Apparitions	
	Communication avec dieu	
	Intervention de dieu dans sa vie	
Ritualiste	Prière	
	Messe	
	Sacrements	
	Pèlerinage	
Conséquentielle	Pardonner	
	Être honnête	

Source : Glock, C. Y. (1963). Origine et Évolution des groupes religieux, Archives de sociologie des religions, vol. 8, p. 29-38, in Quivy 95

Questionnaire : les types d'items

- Items ouverts
  - Assez difficiles à traiter (cf analyse de documents plus bas)
- Items fermés
  - Oui / Non
  - Échelles de Likert (à 3, 4, 5 cases – *« pas d'accord du tout » à « tout à fait d'accord »)*
  - Biodatas (développées plus bas)
  - Échelles continues (curseur), classement de divers éléments par ordre d'importance…
- À questions différentes, données (et tests statistiques) différents
  - Quelle est votre taille ? => donnée métrique
  - Êtes-vous favorable, défavorable ou indifférent à ..? => donnée ordinale
  - Quelle est la marque de votre voiture ? => donnée catégorielle

1) Les questions à se poser
2) La conception du questionnaire
3) La présentation du questionnaire
4) La formulation des questions
5) Le dépouillement de l'enquête



Pour en savoir plus sur la préparation du questionnaire : TD de Vincent GODARD
http://margaux.ipt.univ-paris8.fr/vgodard/enseigne/enquete/memoenq/metd1enq.htm

Optimiser le coût de collecte : l’échantillonnage

Recensement ou sondage ?

- Le Recensement est exhaustif alors que le sondage utilise la technique de l’échantillonnage = sélectionner les sources d’information de manière à collecter des données représentatives de l’ensemble de la situation observée

- Sondage
  - Par rapport à la population totale, interroger suffisamment d’individus pour avoir une marge d’erreur connue … et acceptable.
  - En effet, il existe un phénomène de saturation : après un seuil donné, la collecte d’informations supplémentaires n’apporte que peu de précisions supplémentaires...

Utilisation ou copie interdites sans citation

Plans d'échantillonnage de base

- À l'aveuglette, par volontariat, par jugement d'expert…
- Aléatoire (chaque unité a une chance égale de figurer dans l'échantillon)
- Systématique (choix avec un « pas » de sondage défini)

Source : cours de Jacques Lemaire
(lien cassé) http://www.er.uqam.ca/nobel/k33754/Echantillonnage.pdf

1) Le sondage, un palliatif du recensement ?
2) L'échantillon et sa représentativité
3) Les différentes erreurs
- les erreurs de mesure ;
- l'erreur aléatoire.
- Les non réponses
4) Les différentes méthodes d'échantillonnage



Voir aussi : cours de Vincent GODARD
http://margaux.ipt.univ-paris8.fr/vgodard/enseigne/enquete/memoenq/mem22enq.htm

Plans d'échantillonnage : strates et grappes

- Méthode des strates
  - diviser la population étudiée en sous-ensembles homogènes disjoints ("strates") puis réaliser ensuite un sondage indépendant dans chacune de ces strates (par exemple « jeunes enfants » = CP + CE1 + CE2..)
- Méthode des grappes
  - Travailler de manière exhaustive, mais sur des groupes de population rassemblés (tous les habitants d'un immeuble)

Économies dues à la focalisation sur une catégorie à la fois lors de la collecte.

Méthodes des strates = coût de collecte divisé par 5

Utilisation ou copie interdites sans citation

Plans d’échantillonnage : quotas

- Méthode des quotas

➔ *hypothèse : un échantillon qui ressemble à la population pour un caractère que l’on peut facilement connaître, lui ressemble également pour un autre caractère, celui que l’on cherche à étudier.*

  – La méthode implique une bonne connaissance des statistiques de la population étudiée - Pour être retenue comme variable de contrôle, il faut simultanément avoir *une distribution statistique connue*, *être facile d'observation et être fortement corrélée avec la ou les variables étudiées (par exemple PCS/CSP pour les intentions de vote)*

  – On peut appliquer les quotas en *combinant plusieurs caractères*, par exemple sexe + âge + profession du chef de famille.

Utilisation ou copie interdites sans citation cc

Plans d'échantillonnage : combinaisons

- À deux ou plusieurs degrés (par exemple tirage aléatoire d'un échantillon de localités puis choix des individus par la méthode des quotas)
- À plusieurs phases (échantillon préliminaire pour certaines questions, puis sous-échantillons pour renseignements additionnels).
- Travail sur des panels
  - Groupe de répondants que l'on interroge à intervalles réguliers.
  - Utiliser un panel permet d'identifier des processus d'évolution de manière extrêmement fine et requiert des méthodes d'analyse de données particulières.
- Méthode des itinéraires, ….

Utilisation ou copie interdites sans citation

Plans d'échantillonnage : exemple de note méthodologique

Sondage Ifop

L'évaluation des actions de sécurité de la Préfecture de Police et les attentes des Parisiens en matière de sécurité, février 2002

Note méthodologique

Ce document présente les résultats d'une étude réalisée par l'Ifop. Elle respecte fidèlement les principes scientifiques et déontologiques de l'enquête par sondage.

Étude réalisée pour :

La Préfecture de Police

Échantillon :

Échantillon de 1207 personnes, représentatif de la population parisienne âgée de 18 ans et plus.

La représentativité de l'échantillon a été assurée par la méthode des quotas (sexe, âge, profession du chef de famille) après stratification par arrondissement.

Mode de recueil :

Les interviews ont eu lieu par téléphone au domicile des personnes interrogées.

Dates de terrain du 1 au 7 février 2002.

Vous pouvez télécharger le document au format Word :

Partie I : Analyse par l'Ifop des résultats de l'enquête

Partie II : Les résultats du sondage

Intégralité du document

Utilisation ou copie interdites sans citation

Phénomènes qui menacent la validité

- Phénomène de désirabilité sociale
  - vote « Le Pen », vote Obama (« effet Bradley »)
- Effet de succession de questions (« effet de halo »)
  - Poser une question sur *Hiroshima*, suivie d'une question *sur la nécessité de développer l'énergie nucléaire*….

ou

  - Poser une question sur *l'influence de la circulation automobile sur l'effet de serre*, puis *sur la nécessité de développer l'énergie nucléaire*….
- « quel Etat présente le plus une menace pour la paix dans le monde »… ?
  - *en sous-main, on force le répondant à accepter l'idée que **l'existence** d'un état présente en soit en menace)*
- Lien entre opinion et action
  - *"Un sondage sur les intentions de vote n'a aucun sens si l'on n'a pas une idée du **taux de mobilisation**, c'est-à-dire du pourcentage de gens qui sont sûrs d'aller voter, et du **taux de certitude**, c'est-à-dire du pourcentage de ceux qui sont sûrs d'avoir arrêté leur choix",*

Patrick Lehingue, Coups de sonde dans l'océan des sondages (Ed. du Croquant, 2007)

Utilisation ou copie interdites sans citation

Questionnaire : comment formuler les questions



- La formulation de la question joue un rôle majeur :
  - Perception du sens global : Quel est le sens attribué par le répondant au questionnaire ? Quelle est sa motivation pour répondre ? "gratifications" - points de fidélité, chèques cadeaux, loterie ..
  - Compréhension : Comment le répondant interprète-t-il les mots utilisés ?

- Précautions à prendre

  Attention aux mots utilisés

  Sur une même dimension : poser plusieurs questions

    - En positif et en négatif (désirabilité sociale des réponses positives, ou "biais d'acquiescement")
    - Ne pas mettre ces questions à la suite les unes des autres : mélanger

  Tester les questions (par exemple « *méthode de la pensée à voix haute* » sur des testeurs)

Utilisation ou copie interdites sans citation

Une catégorie de questions particulières : les biodatas

Lautenschlager (94) définit les biodatas comme

"des questions factuelles portant sur la vie et le travail, ainsi que des items mettant en jeu les opinions, valeurs, croyances et attitudes dans une perspective historique".

Stricker et Rock (98) ajoutent que

"ces mesures portent sur des faits relatifs à la vie de la personne et non pas sur de l'introspection et des jugements subjectifs".

Source : *MsA F - Kuschnereit, 2000 - http://www.geocities.com/CollegePark/Quad/5504/*

Utilisation ou copie interdites sans citation

Biodatas : les critères de Maël (1/2)

Maël (91) donne 9 indicateurs qui permettent de classer les items et de définir ce qu'est une biodata idéale. La question posée est :

- **Historique** : signifie qu'elle est relative à des évènements passés, réputés plus fiables que les intentions ou la projection dans ces situations hypothétiques.
- **Externe** : se réfère à des évènements ou des actions, par opposition à des pensées. L'utilisation de données externes est réputées réduire les fausses réponses. En effet, face à des caractéristiques internes, les individus sont plus susceptibles de changer inconsciemment leur réponse.
- **Vérifiable** : il s'agit d'un item qui peut être confirmé par une tierce partie. La vérification elle-même n'est pas indispensable, le fait qu'elle soit possible suffit.
- **Objective** : ce type d'information ne fait appel qu'à la mémoire ("quelle est votre moyenne ?"), par opposition à une appréciation ("êtes-vous bon en maths ?").
- **De première main** : ce type de question se limite à l'expérience personnelle du répondant, par opposition à la nécessité de se mettre à la place d'une autre personne ("que pense X de vous ?").

Utilisation ou copie interdites sans citation

Biodatas : les critères de Maël (2/2)

- **Discrète** : qui mesure un comportement précis (pe un jour donné), par opposition à des comportements agrégés ("d'habitude").
- **Contrôlable par l'individu** : Maël affirme qu'il est contraire à l'éthique de poser des questions sur des évènements sur lesquels le répondant ne peut avoir de contrôle (pe le nombre de frères et sœurs). De plus certains questions peuvent poser des problèmes si elles portent sur des activités d'accès restreint ("jouez-vous au golf ?").
- **Relative au domaine concerné** : pe pour les questionnaires d'embauche les questions doivent se limiter au domaine du travail. Maël distingue les "signes" qui sont des inférences extrapolées à partir d'une attitude ou d'une capacité des "échantillons" qui sont des exemples de comportements. Les échantillons sont censés être de meilleurs prédicteurs de la performance.
- **Non-invasive** : les questions portant sur des sujets comme les pratiques sexuelles, la religion etc… peuvent induire des résistances et la volonté de tromper le questionneur. De plus, elles sont parfois illégales.

 Utilisation ou copie interdites sans citation

Questionnaire : les modes d'administration

Une campagne d'administration est coûteuse … et doit satisfaire aux objectifs imposés par l'échantillonnage

- Comment accéder au répondant ?
  - Aller le chercher (aller chez lui, l'appeler au téléphone)
  - Profiter de son passage à un point donné

- Collecte automatique ou humaine ?
  - Contact humain direct : Porte-à-porte, Sondeurs dans la rue (attention à leur formation / motivation : impact fort)
  - Contact humain indirect : centre d'appel (établir un protocole précis, acheter un fichier) : CATI (Computer-Assisted Telephone Interviewing)
  - Auto-administration : questionnaire papier à rendre ou renvoyer ou CAPI (.. Personal Interviewing), CAWI (.. Web Interviewing)

Utilisation ou copie interdites sans citation

Outils de CAWI (Computer-Assisted Web Interviewing)

- Le spam de mél
  - P.e. avec un questionnaire .doc en PJ
    - Très rapide à lancer, mais
    - Lourd à traiter : réponse à re-saisir
    - ..et non standardisées
    - Pas agréable de répondre
- Sphinx Online : un système « pro »
- Avec Google documents, on peut créer et diffuser un questionnaire en quelques minutes
  - Ouvrir la consultation des résultats aux répondants
  - Changer les questions « à la volée »
  - Gratuit
- …

Source des images : http://www.lesphinx-developpement.fr/ et ce blog

Questionnaire : le contrôle de validité

À entreprendre avant toute interprétation

- Feedback et bilan sur le déroulement de l’enquête
  - Bilan avec les enquêteurs, les répondants..
  - Contrôle de phénomènes parasites
    - « effet enquêteur »
    - Répondant cochant systématiquement la première case…

- Les différentes questions portant sur une même dimension sont-elles corrélées ?
  - Ordre de grandeur : r > .30
  - Remonter aux causes, itérer sur une reconception

Utilisation ou copie interdites sans citation

Questionnaire : le dépouillement

Interprétation et analyse des résultats, analyse

- de la fréquence, de la moyenne, et des moyennes croisées
- de la variance et de l'écart type
- des coefficients de corrélation (régression)
- log-linéaire
- factorielle des correspondances
- en composantes principales
- et interprétation du test d'ajustement du chi-deux
- lexicale pour (questions ouvertes)

Développées dans les cours à venir

Utilisation ou copie interdites sans citation

Questionnaire : la diffusion et l'explication des résultats

En fait, c'est la phase décisive de toute enquête/recherche, mais elle est souvent sous-estimée ou négligée

- Diffusion des réponses brutes aux questions sans explications de la théorie et du modèle sous-jacent
  - Ici : Une bonne façon d'introduire une enquête : présentation des enjeux, questions-réponses argumentées
- Limites de validité non expliquées
- Enquête non contextualisée : mode de collecte mystérieux, un événement important survenu avant l'enquête n'est pas rappelé, questionnaire complet non fourni au lecteur, exploitation de questions isolées
- Présentation des résultats défaillante
  - trop simples (pourcentages),
  - trop complexes (notions incompréhensibles pour le lecteur, pas d'ordres de grandeurs donnés)
  - trompeurs (mapping séduisant de l'ACP).
- Non-contrôle de la diffusion des résultats

L'objectif est de permettre au lecteur/client de se construire son opinion en pleine connaissance de cause.

Utilisation ou copie interdites sans citation

En conclusion : les risques à chaque étape (1/2)

1. Définition de l'objet de l'enquête et de la population cible
   - Le demandeur ne sait pas ce qu'il veut
   - Caractérisation de la population cible incorrecte (trop large, trop restrictive)
   - Ambitions trop larges par rapport aux moyens disponibles
2. Pré-enquête d'exploration du sujet
   - Pas d'exploration ou exploration baclée (pas de contact avec les bons experts, 'cobayes' non représentatifs)
   - Partialité
3. Définition finale des objectifs et hypothèses de l'enquête
   - Hypothèses non explicitées, objectif incorrect
4. Mise au point du questionnaire
   - Mauvaise rédaction des questions ou des explications accompagnant le questionnaire, échelles incorrectes, biais de présentation …

Utilisation ou copie interdites sans citation

En conclusion : les risques à chaque étape (2/2)

5. Test du questionnaire
    - Test trop sommaire (préférer la présence d'un enquêteur pour suivre et noter les réactions des répondants en temps réel)
    - Pas de suivi des préconisations (questions inutiles non supprimées, questions manquantes non ajoutées..)
6. Campagne de collecte de données
    - Échantillon insuffisant ; campagne trop courte
    - Protocole non respecté (typiquement : formation des enquêteurs insuffisante)
7. Bouclage de la campagne de collecte
    - Pas d'analyse des non-réponses, ou des biais de l'échantillon
    - Données finales non validées / stockées sous forme exploitable
    - Contexte de la collecte non mémorisé
8. Dépouillement/analyse des résultats
    - Partialité de l'analyste
    - Méthodologies statistiques pas au point
9. Rapport final
    - Pas de théorie testée ou théorie/ intérêt de l'étude non expliqué
    - Attention aux informations divulguées / confidentialité
    - Ne présente pas les limites de l'enquête, ses méthodes
    - Conclusions peu claires

Utilisation ou copie interdites sans citation

Questions ?

Utilisation ou copie interdites sans citation

Plan

1. Explorer ou vérifier ? Deux catégories d’approches
2. Éventails des démarches de recueil de données
3. Conception de questionnaires
4. Techniques d’entretien et reformulation
5. L'Analyse Factorielle des Correspondances pour les nuls
6. Validité et Fiabilité des données

Utilisation ou copie interdites sans citation



Rémi BACHELET

Enseignant-chercheur,
Ecole Centrale de Lille

Mon CV est disponible ici.

Mes principaux cours à Centrale
Gestion de projet, sociologie des organisations, recueil, analyse et traitement de données, prévention du plagiat, module de marchés financiers, cours de qualité et méthodes de résolution de problèmes, établir des cartes conceptuelles, utiliser Wikipédia et CentraleWiki, formation au coaching pédagogique et à l'encadrement

Utilisation ou copie interdites sans citation